“Too young to die”. Deprivation measures combining poverty and premature mortality. ONLINE APPENDIX

Jean-Marie Baland[*], Guilhem Cassan[†], Benoit Decerf[‡]

July 28, 2020

1 Non-monotonic evolution of GD

Consider a stationary population with one individual born every year who lives exactly for 4 periods, with a mortality rate at age 3 equal to 1. The mortality vector is thus $\mu = (0, 0, 0, 1, \dots)$. We assume that the age threshold, $\hat{a}$, is equal to 12, and $\gamma = 1$. There is no alive deprivation. GD for this situation is equal to 8/12, and is equal to ID. In period t^s, there is a permanent mortality shock such that the new mortality rate at age 1 is equal to 1. The new mortality vector is thus $\mu^s = (0, 1, 0, 1, \dots)$. Table 1 summarizes the evolution of this population after this permanent shock.

Table 1: Non-monotonicity of GD indices after permanent mortality shock.

period	0	1	2	3	4	...	11	GD
$t < t^s$	NP	NP	NP	NP	D	...	D	$8/12 = 0.66$
t^s	NP	NP	NP	NP	D	...	D	$18/22 = 0.82$
$t^s + 1$	NP	NP	D	NP	D	...	D	$18/21 = 0.86$
$t^s + 2$	NP	NP	D	D	D	...	D	$10/12 = 0.83$

Two individuals die at the end of period t^s and GD records 18 PYPLs. Given that four individuals lived in period t^s, GD is equal to 18/22. In period $t^s + 1$, there is no

[*]University of Namur.
[†]University of Namur.
[‡]University of Namur. benoit.decerf@unamur.be

individual of age 2, and one individual of age 0, 1 and 3. GD records again 18 PYPLs, but given that only three individuals were alive, GD is equal to 18/21. Two periods after the shock, the new demographic equilibrium is such that there are only two individuals alive, of age 1 and 2 respectively. There are 10 PYPLs, out of a total of 12, so that GD is equal to 10/12. Because of the mechanical adaptation of the population pyramid, GD increases in $t^s + 1$ but decreases in $t^s + 2$.

How should we think about the non-monotonic behavior of GD? This behavior reflects the evolution of the population. Indeed, the presence of the 3-years old individual in $t^s + 1$ implies that the mortality vector μ^s does create more PYPLs in $t^s + 1$ than in $t^s + 2$. So GD conveys correct information about actual deprivation. However, a fixed mortality vector μ^s is related to the fundamentals for a population's health situation. One should therefore not necessarily conclude from the evolution of GD that these fundamentals have necessarily changed.

2 Proofs

2.1 Proof of Proposition 1

It is easy to check that ID satisfies the seven axioms, so that the proof of necessity is omitted. Herebelow, we concentrate on the proof of sufficiency.

Let $\mathcal{Q}_+$ denote the set of non-negative rational numbers. Consider Δ, the 2-simplex on rational numbers, i.e. $\Delta = \{\mathbf{v} \in \mathcal{Q}_+^3 \mid \mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3 = 1\}$.

Step 1: Construct a mapping $m : \mathcal{X} \to \Delta$ such that $m(\mathcal{X}) = \Delta$ and for any two $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$, if $m(\mathbf{x}) = m(\mathbf{x}')$ then $P(\mathbf{x}) = P(\mathbf{x}')$.

We construct mapping m as the composition of four mappings, i.e. $m(\mathbf{x}) = m^4 \circ m^3 \circ m^2 \circ m^1(\mathbf{x})$.[1]

First, mapping m^1 removes all individuals who are not in the reference population. Let $\mathcal{X}^*$ be the subset of distributions that do not have any individual who is dead and was born at least $\hat{a}$ years before t, i.e. $\mathcal{X}^* = \{\mathbf{x} \in \mathcal{X} \mid b_i > t - \hat{a} \text{ for all } i \text{ for whom } s_i = D\}$. Let mapping $m^1 : \mathcal{X} \to \mathcal{X}^*$ return for any $\mathbf{x} \in \mathcal{X}$ the image $\mathbf{x}^* = m^1(\mathbf{x})$ with $n(\mathbf{x}^*) = f(\mathbf{x}) + p(\mathbf{x}) + d(\mathbf{x})$ and for any $i \leq n(\mathbf{x}^*)$ the i^{th} component of $\mathbf{x}^*$ is defined as $\mathbf{x}_i^* \equiv \mathbf{x}_j$, where j is the i^{th} individual in $\mathbf{x}$ for whom either $s_i \neq D$ or $s_i = D$ and $b_i > t - \hat{a}$. By the definition of mapping m^1, we have for all $\mathbf{x} \in \mathcal{X}^*$ that $m^1(\mathbf{x}) = \mathbf{x}$.

[1]The composite mapping m is defined a $m(\mathbf{x}) = m^4(m^3(m^2(m^1(\mathbf{x}))))$.

Hence, $m^1(\mathcal{X}) = \mathcal{X}^*$. Also, any two $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$ for which $m^1(\mathbf{x}) = m^1(\mathbf{x}')$ are such that $P(\mathbf{x}) = P(\mathbf{x}')$ by Weak Independence of Dead and Anonymity.

Second, mapping m^2 removes the birth year of all remaining individuals. Recalling that $\mathcal{S} = \{NP, AP, D\}$, let mapping $m^2 : \mathcal{X}^* \to \cup_{n\in\mathcal{N}}\mathcal{S}^n$ return for any $\mathbf{x} \in \mathcal{X}^*$ the image $\mathbf{o} = m^2(\mathbf{x})$ with $n(\mathbf{o}) = n(\mathbf{x})$ and for any $i \leq n(\mathbf{o})$ the i^{th} component of $\mathbf{o}$ is defined from $\mathbf{x}_i = (b_i, s_i)$ as $\mathbf{o}_i \equiv s_i$. By construction of m^2, we have $m^2(\mathcal{X}^*) = \cup_{n\in\mathcal{N}}\mathcal{S}^n$. By construction of m^1, all dead individuals in a distribution $\mathbf{x} \in \mathcal{X}^*$ are *prematurely* dead. Hence, if $\mathbf{o}_i = D$, then i is *prematurely* dead. Therefore, any two $\mathbf{x}, \mathbf{x}' \in \mathcal{X}^*$ for which $m^2(\mathbf{x}) = m^2(\mathbf{x}')$ are such that $P(\mathbf{x}) = P(\mathbf{x}')$ by Weak Independence of Birth Year.

Third, mapping m^3 counts the number of individuals exhibiting each status. Consider the set $\mathcal{N}_0^3\backslash_{(0,0,0)}$, which contains all triplets of numbers in $\mathcal{N}_0 = \{0, 1, 2, \dots\}$ except the nul triplet $(0, 0, 0)$. Let the mapping $m^3 : \cup_{n\in\mathcal{N}}\mathcal{S}^n \to \mathcal{N}_0^3\backslash_{(0,0,0)}$ return for any $\mathbf{o} \in \cup_{n\in\mathcal{N}}\mathcal{S}^n$ the image $\mathbf{w} = m^3(\mathbf{o})$ such that $\mathbf{w}_1 \equiv \#\{i \leq n(\mathbf{o}) \mid \mathbf{o}_i = NP\}$, $\mathbf{w}_2 \equiv \#\{i \leq n(\mathbf{o}) \mid \mathbf{o}_i = AP\}$ and $\mathbf{w}_3 \equiv \#\{i \leq n(\mathbf{o}) \mid \mathbf{o}_i = D\}$.[2] By construction, we have $m^3 \circ m^2(\mathcal{X}^*) = \mathcal{N}_0^3\backslash_{(0,0,0)}$.[3] Also, any two $\mathbf{x}, \mathbf{x}' \in \mathcal{X}^*$ for which $m^3 \circ m^2(\mathbf{x}) = m^3 \circ m^2(\mathbf{x}')$ are such that $P(\mathbf{x}) = P(\mathbf{x}')$ by Anonymity and Weak Independence of Birth Year.

Fourth, mapping m^4 computes the fraction of individuals having each status. Let mapping $m^4 : \mathcal{N}_0^3\backslash_{(0,0,0)} \to \Delta$ return for any $\mathbf{w} \in \mathcal{N}_0^3\backslash_{(0,0,0)}$ the image $v = m^4(\mathbf{w})$ defined as

$$\mathbf{v} = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) \equiv \left(\frac{\mathbf{w}_1}{\mathbf{w}_1 + \mathbf{w}_2 + \mathbf{w}_3}, \frac{\mathbf{w}_2}{\mathbf{w}_1 + \mathbf{w}_2 + \mathbf{w}_3}, \frac{\mathbf{w}_3}{\mathbf{w}_1 + \mathbf{w}_2 + \mathbf{w}_3} \right),$$

where $\mathbf{v}_1$ is the fraction of non-poor, $\mathbf{v}_2$ is the fraction of poor and $\mathbf{v}_3$ is the fraction of prematurely dead. Let mapping $m : \mathcal{X} \to \Delta$ be defined as $m(\mathbf{x}) = m^4 \circ m^3 \circ m^2 \circ m^1(\mathbf{x})$.

First, we show that for any $\mathbf{v} \in \Delta$ there exists a $\mathbf{x} \in \mathcal{X}^*$ such that $m(\mathbf{x}) = \mathbf{v}$. As $\mathbf{v} \in \Delta$, there exist $c_1, c_2, c_3, e_1, e_2, e_3 \in \mathcal{N}$ such that $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = (c_1/e_1, c_2/e_2, c_3/e_3)$. Consider any distribution $\mathbf{x}$ with $n(\mathbf{x}) = e_1e_2e_3$, where $c_1e_2e_3$ individuals are non-poor, $c_2e_1e_3$ individuals are poor, and $c_3e_1e_2$ individuals are prematurely dead. As $\mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3 = 1$, we have that $c_1e_2e_3 + c_2e_1e_3 + c_3e_1e_2 = e_1e_2e_3$. All individuals in $\mathbf{x}$ who are dead are prematurely dead, hence, $\mathbf{x} \in \mathcal{X}^*$. By construction of $\mathbf{x}$, we have $m(\mathbf{x}) = \mathbf{v}$.

There remains to show that for any two $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$ such that $m(\mathbf{x}) = m(\mathbf{x}')$ we have $P(\mathbf{x}) = P(\mathbf{x}')$. We have shown above that if $m^3 \circ m^2 \circ m^1(\mathbf{x}) = m^3 \circ m^2 \circ m^1(\mathbf{x}')$, then $P(\mathbf{x}) = P(\mathbf{x}')$. There remains to show that if $m^3 \circ m^2 \circ m^1(\mathbf{x}) \neq m^3 \circ m^2 \circ$

[2] For any set A, we denote the cardinality of A by $\#A$.
[3] By definition of $\mathcal{X}$, there is no $\mathbf{x} \in \mathcal{X}^*$ such that $m^3 \circ m^2(\mathbf{x}) = (0, 0, 0)$.

$m^1(\mathbf{x}')$ and $m(\mathbf{x}) = m(\mathbf{x}')$, we have $P(\mathbf{x}) = P(\mathbf{x}')$. To do so, we show that for any two $\mathbf{w}, \mathbf{w}' \in \mathcal{N}_0^3 \backslash_{(0,0,0)}$ such that $m^4(\mathbf{w}) = m^4(\mathbf{w}')$, there exist $\mathbf{y}, \mathbf{y}' \in \mathcal{X}$ such that $m^3 \circ m^2 \circ m^1(\mathbf{y}) = \mathbf{w}$, $m^3 \circ m^2 \circ m^1(\mathbf{y}') = \mathbf{w}'$ and $P(\mathbf{y}) = P(\mathbf{y}')$. By construction of mapping m^4, any two $\mathbf{w}, \mathbf{w}' \in \mathcal{N}_0^3 \backslash_{(0,0,0)}$ for which $m^4(\mathbf{w}) = m^4(\mathbf{w}')$ are such that for $k = \mathbf{w}_1' + \mathbf{w}_2' + \mathbf{w}_3'$ and $k' = \mathbf{w}_1 + \mathbf{w}_2 + \mathbf{w}_3$ we have a $\mathbf{w}'' \in \mathcal{N}_0^3 \backslash_{(0,0,0)}$ such that $\mathbf{w}'' = k\mathbf{w} = k'\mathbf{w}'$. Then, there exist $\mathbf{y}, \mathbf{y}', \mathbf{y}'', \mathbf{y}''' \in \mathcal{X}^*$ with $m^3 \circ m^2 \circ m^1(\mathbf{y}) = \mathbf{w}$, $m^3 \circ m^2 \circ m^1(\mathbf{y}') = \mathbf{w}'$, $m^3 \circ m^2 \circ m^1(\mathbf{y}'') = m^3 \circ m^2 \circ m^1(\mathbf{y}''') = \mathbf{w}''$ such that $\mathbf{y}''$ is a k-replication of $\mathbf{y}$ and $\mathbf{y}'''$ is a k'-replication of $\mathbf{y}'$. By Replication Invariance, we have that $P(\mathbf{y}) = P(\mathbf{y}'')$ and $P(\mathbf{y}') = P(\mathbf{y}''')$. As $m^3 \circ m^2 \circ m^1(\mathbf{y}'') = m^3 \circ m^2 \circ m^1(\mathbf{y}''')$, we have $P(\mathbf{y}'') = P(\mathbf{y}''')$. Together, $P(\mathbf{y}) = P(\mathbf{y}')$.

Step 2: Using mapping m, define an ordering $\succeq$ on Δ from the ordering on $\mathcal{X}$ represented by P.[4]

Let $\succeq$ be an ordering on Δ defined such that for any two $\mathbf{v}, \mathbf{v}' \in \Delta$ we have $\mathbf{v} \succ \mathbf{v}'$ (resp. $\mathbf{v} \sim \mathbf{v}'$) if there exist $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$ such that $\mathbf{v} = m(\mathbf{x})$ and $\mathbf{v}' = m(\mathbf{x}')$ and $P(\mathbf{x}) < P(\mathbf{x}')$ (resp. $P(\mathbf{x}) = P(\mathbf{x}')$). We showed at the end of Step 1 that there always exist $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$ such that $\mathbf{v} = m(\mathbf{x})$ and $\mathbf{v}' = m(\mathbf{x}')$, which shows that ordering $\succeq$ is complete. Moreover, any two $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$ with $m(\mathbf{x}) = m(\mathbf{x}')$ are such that $P(\mathbf{x}) = P(\mathbf{x}')$, which shows that ordering $\succeq$ is well-defined. Together, we have that for any two $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$ and $\mathbf{v}, \mathbf{v}' \in \Delta$ with $\mathbf{v} = m(\mathbf{x})$ and $\mathbf{v}' = m(\mathbf{x}')$, we have

$$P(\mathbf{x}) \leq P(\mathbf{x}') \Leftrightarrow \mathbf{v} \succeq \mathbf{v}'. \tag{1}$$

Step 3: Identifying the appropriate value for γ.

First, we show that $\succeq$ satisfies the following **convexity property**: for any two $\mathbf{v}, \mathbf{v}' \in \Delta$ with $\mathbf{v} \succ \mathbf{v}'$ and any rational $\lambda \in (0, 1)$ we have $\mathbf{v} \succ \lambda\mathbf{v} + (1-\lambda)\mathbf{v}' \succ \mathbf{v}'$. Take any two $\mathbf{x}, \mathbf{y} \in \mathcal{X}^*$ such that $\mathbf{v} = m(\mathbf{x})$ and $\mathbf{v}' = m(\mathbf{y})$. Using Replication Invariance, these two distributions can be taken such that $n(\mathbf{x}) = n(\mathbf{y})$, which we assume henceforth. By Equivalence (1), we have $P(\mathbf{x}) < P(\mathbf{y})$. By definition of λ, there exists $c, e \in \mathcal{N}$ such that $\lambda = c/e$. Let $\mathbf{x}^c$ be a c-replication of $\mathbf{x}$, $\mathbf{x}^{(e-c)}$ be a $(e-c)$-replication of $\mathbf{x}$, $\mathbf{y}^c$ be a c-replication of $\mathbf{y}$ and $\mathbf{y}^{(e-c)}$ be a $(e-c)$-replication of $\mathbf{y}$. By construction, we have

[4] An ordering is a complete, reflexive and transitive binary relation.

$n(\mathbf{x}^c) = n(\mathbf{y}^c)$ and $n(\mathbf{x}^{(e-c)}) = n(\mathbf{y}^{(e-c)})$. By Replication Invariance, we have

$$P(\mathbf{x}^c) = P(\mathbf{x}^{(e-c)}) = P(\mathbf{x}^c, \mathbf{x}^{(e-c)}) < P(\mathbf{y}^c) = P(\mathbf{y}^{(e-c)}) = P(\mathbf{y}^c, \mathbf{y}^{(e-c)}).$$

As all these distributions belong to $\mathcal{X}^*$, we have by Subgroup Consistency that $P(\mathbf{x}^c, \mathbf{x}^{(e-c)}) < P(\mathbf{x}^c, \mathbf{y}^{(e-c)})$ and that $P(\mathbf{x}^c, \mathbf{y}^{(e-c)}) < P(\mathbf{y}^c, \mathbf{y}^{(e-c)})$. Now, we constructed these replications such that $\mathbf{v} = m(\mathbf{x}^c, \mathbf{x}^{(e-c)})$, $\mathbf{v}' = m(\mathbf{y}^c, \mathbf{y}^{(e-c)})$ and also $\lambda\mathbf{v} + (1-\lambda)\mathbf{v}' = m(\mathbf{x}^c, \mathbf{y}^{(e-c)})$. This yields the desired result by Equivalence (1).

Second, we derive the value $\gamma > 0$ for which P is ordinally equivalent to ID_γ. Let the three vertices $(1,0,0), (0,1,0), (0,0,1) \in \Delta$ be respectively denoted by $\mathbf{v}^{100}$, $\mathbf{v}^{010}$ and $\mathbf{v}^{001}$. By Least Deprivation and Equivalence (1), we have that $\mathbf{v}^{100} \succ \mathbf{v}^{010}$ and $\mathbf{v}^{100} \succ \mathbf{v}^{001}$. There are three cases.

- **Case 1: $\mathbf{v}^{010} \sim \mathbf{v}^{001}$.**

  Take $\gamma = 1$.

- **Case 2: $\mathbf{v}^{010} \succ \mathbf{v}^{001}$.**

  Consider the edge connecting vertices $\mathbf{v}^{100}$ and $\mathbf{v}^{001}$, which we denote by $E^{100}_{001} = \{\mathbf{v} \in \Delta \mid \mathbf{v}_2 = 0\}$. As $\mathbf{v}^{100} \succ \mathbf{v}^{001}$, the convexity property implies that for any $\mathbf{v}, \mathbf{v}' \in E^{100}_{001}$, if $\mathbf{v}_1 > \mathbf{v}'_1$ then $\mathbf{v} \succ \mathbf{v}'$ and if $\mathbf{v}_1 < \mathbf{v}'_1$ then $\mathbf{v} \prec \mathbf{v}'$. Let $\Delta^{\mathcal{R}_+}$ be the 2-simplex on the set of real numbers. As $\mathbf{v}^{100} \succ \mathbf{v}^{010} \succ \mathbf{v}^{001}$, there exists a $\mathbf{v}^* \in \Delta^{\mathcal{R}_+}$ on the edge connecting the two vertices $\mathbf{v}^{100}$ and $\mathbf{v}^{001}$ such that for any $\mathbf{v} \in E^{100}_{001}$, if $\mathbf{v}_1 > \mathbf{v}^*_1$ then $\mathbf{v} \succ \mathbf{v}^{010}$ and, if $\mathbf{v}_1 < \mathbf{v}^*_1$ then $\mathbf{v} \prec \mathbf{v}^{010}$. Moreover, if $\mathbf{v}^* \in \Delta$, then $\mathbf{v}^* \sim \mathbf{v}^{010}$ (see proof below). As $\mathcal{Q}$ is dense in $\mathcal{R}$, there is always an irrational between two rationals. Therefore, $\mathbf{v}^*$ is the unique element of $\Delta^{\mathcal{R}_+}$ with these properties.

  We show that if $\mathbf{v}^* \in \Delta$, then $\mathbf{v}^* \sim \mathbf{v}^{010}$. Consider the contradiction assumption that $\mathbf{v}^* \in \Delta$ and $\mathbf{v}^* \succ \mathbf{v}^{010}$.[5] We construct a $\mathbf{v}' \in E^{100}_{001}$ such that $\mathbf{v}'_1 < \mathbf{v}^*_1$ and $\mathbf{v}' \succ \mathbf{v}^{010}$. Such $\mathbf{v}'$ is in contradiction with the definition of $\mathbf{v}^*$, which requires that for any $\mathbf{v}' \in E^{100}_{001}$ with $\mathbf{v}'_1 < \mathbf{v}^*_1$ we have $\mathbf{v}' \prec \mathbf{v}^{010}$. We construct $\mathbf{v}' \in E^{100}_{001}$ as follows. Take any two distributions $\mathbf{x}, \mathbf{y} \in \mathcal{X}^*$ such that $\mathbf{v}^{010} = m(\mathbf{x})$ and $\mathbf{v}^* = m(\mathbf{y})$. As $\mathbf{v}^* \succ \mathbf{v}^{010}$, we have by Equivalence (1) that $P(\mathbf{x}) > P(\mathbf{y})$. Let $\mathbf{z} \in \mathcal{X}^*$ be a distribution with $n(\mathbf{z}) = 1$ and whose unique individual is prematurely dead. By Young Continuity, there exists some k such that $P(\mathbf{x}) > P(\mathbf{y}^k, \mathbf{z})$. Consider

[5] The alternative contradiction assumption for which $\mathbf{v}^* \in \Delta$ and $\mathbf{v}^* \prec \mathbf{v}^{010}$ also leads to an impossibility.

$\mathbf{v}' = m(\mathbf{y}^k, \mathbf{z})$. By Equivalence (1), we have $\mathbf{v}' \succ \mathbf{v}^{010}$. As $\mathbf{v}^* \in E_{001}^{100}$, we have by construction that $\mathbf{v}' \in E_{001}^{100}$ and $\mathbf{v}'_1 < \mathbf{v}^*_1$, the desired result.

We take $\gamma = \frac{1}{\mathbf{v}^*_3}$. We have $\mathbf{v}^*_3 \in (0,1)$ because $\mathbf{v}^{010} \succ \mathbf{v}^{001}$ and $\mathbf{v}^{100} \succ \mathbf{v}^{010}$ respectively imply that $\mathbf{v}^* \neq \mathbf{v}^{001}$ and $\mathbf{v}^* \neq \mathbf{v}^{100}$. As $\mathbf{v}^*_3 \in (0,1)$, we have $\gamma > 1$.

- **Case 3:** $\mathbf{v}^{010} \prec \mathbf{v}^{001}$.

  The construction of γ is similar to that proposed in Case 2. We find the unique element $\mathbf{v}^{**} \in \Delta^{\mathcal{R}_+}$ that splits the edge from $\mathbf{v}^{100}$ to $\mathbf{v}^{010}$ between elements $\mathbf{v}$ for which $\mathbf{v} \succ \mathbf{v}^{001}$ and elements $\mathbf{v}'$ for which $\mathbf{v}' \prec \mathbf{v}^{001}$. We take $\gamma = \mathbf{v}^{**}_2$ and have $\gamma \in (0,1)$.

We assume henceforth that Case 2 applies, i.e. $\mathbf{v}^{010} \succ \mathbf{v}^{001}$. We omit the proof for Case 1 that is simpler and the proof for Case 3 that is very similar.

Step 4: Show that ID_γ is ordinally equivalent to P.

Let function $F : \Delta^{\mathcal{R}_+} \to \mathcal{R}_-$ be defined by $F(\mathbf{v}) = -(\mathbf{v}_2 + \gamma \mathbf{v}_3)$. By construction of mapping m and the definition of ID_γ, for any $\mathbf{v} \in \Delta$ and any $\mathbf{x} \in \mathcal{X}$ such that $\mathbf{v} = m(\mathbf{x})$ we have that $F(\mathbf{v}) = -ID_\gamma(\mathbf{x})$. If we show that F represents the ordering $\succeq$ on Δ, then we get from Equivalence (1) that ID_γ is ordinally equivalent to P, the desired result.

First, we show that for any $\mathbf{v} \in \Delta$ we have $\mathbf{v} \succeq \mathbf{v}^{010}$ if and only if $F(\mathbf{v}) \geq F(\mathbf{v}^{010})$. By definition of F, we have that $F(\mathbf{v}^{100}) = 0$, $F(\mathbf{v}^{010}) = -1$ and $F(\mathbf{v}^{001}) = -\gamma$. Partition Δ into three subsets, i.e. $\Delta = \Delta^{100} \cup \Delta^{010} \cup \Delta^{001}$ defined as $\Delta^{010} = \{\mathbf{v} \in \Delta \mid F(\mathbf{v}) = -1\}$, $\Delta^{100} = \{\mathbf{v} \in \Delta \mid F(\mathbf{v}) > -1\}$ and $\Delta^{001} = \{\mathbf{v} \in \Delta \mid F(\mathbf{v}) < -1\}$.[6] We need to show that any $\mathbf{v} \in \Delta^{100}$ is such that $\mathbf{v} \succ \mathbf{v}^{010}$, any $\mathbf{v} \in \Delta^{010}$ is such that $\mathbf{v} \sim \mathbf{v}^{010}$ and any $\mathbf{v} \in \Delta^{001}$ is such that $\mathbf{v} \prec \mathbf{v}^{010}$. In order to avoid repetitions, we only prove that any $\mathbf{v} \in \Delta^{100}$ is such that $\mathbf{v} \succ \mathbf{v}^{010}$. To do so, we show that $\mathbf{v} = \lambda \mathbf{v}^{010} + (1-\lambda)\mathbf{v}'$ for some rational $\lambda \in [0,1)$ and some $\mathbf{v}'$ on the edge E_{001}^{100} with $\mathbf{v}'_1 > \mathbf{v}^*_1$. This construction is illustrated in Panel A of Figure 1. Given that any $\mathbf{v}' \in E_{001}^{100}$ for which $\mathbf{v}'_1 > \mathbf{v}^*_1$ is such that $\mathbf{v}' \succ \mathbf{v}^{010}$, the convexity property of $\succeq$ then implies that $\mathbf{v} \succ \mathbf{v}^{010}$. Take $\mathbf{v}' \equiv \left(1 - \frac{\mathbf{v}_3}{1-\mathbf{v}_2}, 0, \frac{\mathbf{v}_3}{1-\mathbf{v}_2}\right)$. As $\mathbf{v} \in \Delta$, the definition of $\mathbf{v}'$ is such that $\mathbf{v}' \in E_{001}^{100}$. Let $\mathbf{v}'' \equiv \lambda \mathbf{v}^{010} + (1-\lambda)\mathbf{v}'$ where $\lambda \equiv \mathbf{v}_2 \in [0,1)$ since $\mathbf{v} \in \Delta^{100}$. We have $\mathbf{v}'' = \mathbf{v}$ since, by

[6] We have defined F and γ such that $F(\mathbf{v}^*) = F(\mathbf{v}^{010})$. From a geometric perspective, the set of elements $\mathbf{v} \in \Delta^{\mathcal{R}_+}$ for which $F(\mathbf{v}) = -1$ is the segment connecting $\mathbf{v}^{010}$ with $\mathbf{v}^*$. Observe that if $\mathbf{v}^* \notin \Delta$, then the only element in this segment belonging to Δ is the vertex $\mathbf{v}^{010}$ and, therefore, Δ^{010} degenerates to $\{\mathbf{v}^{010}\}$. The subset Δ^{100} contains vertex $\mathbf{v}^{100}$ and all elements of Δ that are on $\mathbf{v}^{100}$'s side of the segment connecting $\mathbf{v}^{010}$ with $\mathbf{v}^*$. In turn, Δ^{001} contains vertex $\mathbf{v}^{001}$ and all elements of Δ that are on $\mathbf{v}^{001}$'s side of the segment connecting $\mathbf{v}^{010}$ with $\mathbf{v}^*$.

construction of $\mathbf{v}'$, we have $\mathbf{v}''_2 = \lambda = \mathbf{v}_2$ and $\mathbf{v}''_3 = (1-\lambda)\mathbf{v}'_3 = \mathbf{v}_3$. There remains to show that $\mathbf{v}'_1 > \mathbf{v}^*_1$. Last inequality is equivalent to $1 - \frac{\mathbf{v}_3}{1-\mathbf{v}_2} > 1 - \frac{1}{\gamma}$, which simplifies to $\frac{1}{\gamma} > \frac{\mathbf{v}_3}{1-\mathbf{v}_2}$. This inequality holds because, as $\mathbf{v} \in \Delta^{100}$, we have $F(\mathbf{v}) > -1$, which simplifies to the same inequality.

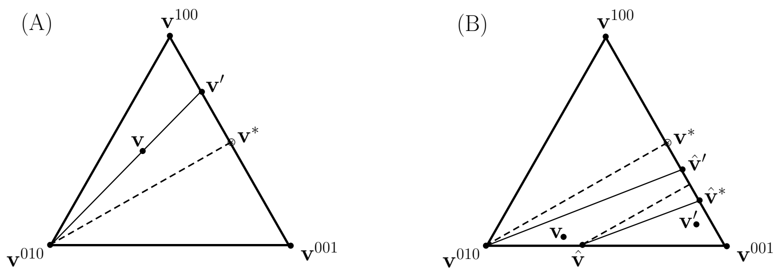

Figure 1: Panel A: construction used in order to show that $\mathbf{v} \succ \mathbf{v}^{010}$ when $F(\mathbf{v}) > F(\mathbf{v}^{010})$. Panel B: construction used in order to show that $\mathbf{v} \succ \mathbf{v}'$ when $F(\mathbf{v}) > F(\mathbf{v}')$. Iso-$F$ lines are dashed.

Geometrically, we have shown that the segment connecting $\mathbf{v}^{010}$ to $\mathbf{v}^*$ is an "implicit" indifference curve of $\succeq$.[7] The intuition for the rest of the proof is that all parallel segments are also "implicit" indifference curves of $\succeq$.

Take any two $\mathbf{v}, \mathbf{v}' \in \Delta$ with $F(\mathbf{v}) \geq F(\mathbf{v}')$, we show that $\mathbf{v} \succeq \mathbf{v}'$. If $F(\mathbf{v}) \geq -1 \geq F(\mathbf{v}')$, then the previous argument directly yields the result. We focus on the particular case $-1 > F(\mathbf{v}) > F(\mathbf{v}')$ and show that $\mathbf{v} \succ \mathbf{v}'$ (the proofs for the other cases are similar). The construction is illustrated in Panel B of Figure 1. This case is such that there exists a $\hat{\mathbf{v}} = (0, \hat{\mathbf{v}}_2, 1 - \hat{\mathbf{v}}_2) \in E^{010}_{001}$ with $F(\mathbf{v}) > F(\hat{s}) > F(\mathbf{v}')$, because $F(\mathbf{v}^{010}) = -1$ and $F(\mathbf{v}^{001}) = \min_{\mathbf{v}'' \in \Delta} F(\mathbf{v}'')$. By the convexity property of $\succeq$, our assumption $\mathbf{v}^{010} \succ \mathbf{v}^{001}$ implies that $\mathbf{v}^{010} \succ \hat{\mathbf{v}} \succeq \mathbf{v}^{001}$. Therefore, there exists a unique $\hat{\mathbf{v}}^* \in \Delta^{\mathcal{R}_+}$ on the edge connecting vertices $\mathbf{v}^{100}$ and $\mathbf{v}^{001}$ such that for any $\mathbf{v}'' \in E^{100}_{001}$, if $\mathbf{v}''_1 > \hat{\mathbf{v}}^*_1$ then $\mathbf{v}'' \succ \hat{\mathbf{v}}$, if $\mathbf{v}''_1 < \hat{\mathbf{v}}^*_1$ then $\mathbf{v}'' \prec \hat{\mathbf{v}}$ and if $\hat{\mathbf{v}}^* \in \Delta$, then $\hat{\mathbf{v}}^* \sim \hat{\mathbf{v}}$ (the omitted proof for this claim follows the argument provided in Step 3 Case 2).

First, we show that the segment connecting $\hat{\mathbf{v}}$ to $\hat{\mathbf{v}}^*$ is parallel to the segment connecting $\mathbf{v}^{010}$ to $\mathbf{v}^*$. Formally, this is equivalent to showing that $\hat{\mathbf{v}}_2 = \frac{\hat{\mathbf{v}}^*_1}{\mathbf{v}^*_1}$. Consider the contradiction assumption for which $\hat{\mathbf{v}}_2 > \frac{\hat{\mathbf{v}}^*_1}{\mathbf{v}^*_1}$.[8] Assume that $\hat{\mathbf{v}}^* \in \Delta$.[9] Consider now

[7] We call this indifference curve "implicit" because it is defined in $\Delta^{\mathcal{R}_+}$ rather than in Δ.

[8] The alternative contradiction assumption for which $\hat{\mathbf{v}}_2 < \frac{\hat{\mathbf{v}}^*_1}{\mathbf{v}^*_1}$ also leads to an impossibility.

[9] If $\hat{\mathbf{v}}^* \notin \Delta$, then replace $\hat{\mathbf{v}}^*$ by a nearby $\tilde{\mathbf{v}}^* \in E^{100}_{001}$ for which $\tilde{\mathbf{v}}^*_1 > \hat{\mathbf{v}}^*_1$ and $\hat{\mathbf{v}}_2 > \frac{\tilde{\mathbf{v}}^*_1}{\mathbf{v}^*_1}$. As $\tilde{\mathbf{v}}^*_1 > \hat{\mathbf{v}}^*_1$, we have $\tilde{\mathbf{v}}^* \succ \hat{\mathbf{v}}$.

$\hat{\mathbf{v}}^{'} = (\frac{\hat{\mathbf{v}}_1^*}{\hat{\mathbf{v}}_2}, 0, 1 - \frac{\hat{\mathbf{v}}_1^*}{\hat{\mathbf{v}}_2}) \in E_{001}^{100}$. By the contradiction assumption, we have $\hat{\mathbf{v}}_1^{'} < \mathbf{v}_1^*$ and, hence, $\mathbf{v}^{010} \succ \hat{\mathbf{v}}^{'}$. By construction, for the rational $\lambda = \hat{\mathbf{v}}_2$ we have:

$$\hat{\mathbf{v}} = \lambda \mathbf{v}^{010} + (1-\lambda)\mathbf{v}^{001} \qquad \text{and} \qquad \hat{\mathbf{v}}^* = \lambda \hat{\mathbf{v}}^{'} + (1-\lambda)\mathbf{v}^{001}.$$

We use the fact that $\mathbf{v}^{010} \succ \hat{\mathbf{v}}^{'}$ in order to show that $\hat{\mathbf{v}} \succ \hat{\mathbf{v}}^*$, a contradiction to the definition of $\hat{\mathbf{v}}^*$. Take any three distributions $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{X}^*$ such that $\mathbf{v}^{010} = m(\mathbf{x})$, $\hat{\mathbf{v}}^{'} = m(\mathbf{y})$ and $\mathbf{v}^{001} = m(\mathbf{z})$. By Equivalence (1), we have $P(\mathbf{x}) < P(\mathbf{y}) < P(\mathbf{z})$. Using Replication Invariance, these three distributions can be taken such that $n(\mathbf{x}) = n(\mathbf{y}) = n(\mathbf{z})$, which we assume henceforth. As $\lambda = \hat{\mathbf{v}}_2$, there exist $c, e \in \mathcal{N}$ such that $\lambda = c/e$. Let $\mathbf{x}^c$ be a c-replication of $\mathbf{x}$, $\mathbf{y}^c$ be a c-replication of $\mathbf{y}$ and $\mathbf{z}^{(e-c)}$ be a $(e-c)$-replication of $\mathbf{z}$. By Replication Invariance, we have $P(\mathbf{x}^c) < P(\mathbf{y}^c) < P(\mathbf{z}^{(e-c)})$. Thus, by Subgroup Consistency, we have that $P(\mathbf{x}^c, \mathbf{z}^{(e-c)}) < P(\mathbf{y}^c, \mathbf{z}^{(e-c)})$. Now, we constructed these replications such that $\hat{\mathbf{v}} = m((\mathbf{x}^c, \mathbf{z}^{(e-c)}))$ and $\hat{\mathbf{v}}^* = m((\mathbf{y}^c, \mathbf{z}^{(e-c)}))$. By Equivalence (1), we obtain $\hat{\mathbf{v}} \succ \hat{\mathbf{v}}^*$, the desired contradiction. Therefore, we have $\hat{\mathbf{v}}_2 = \frac{\hat{\mathbf{v}}_1^*}{\mathbf{v}_1^*}$, which implies that $F(\hat{\mathbf{v}}) = F(\hat{\mathbf{v}}^*)$ as $\gamma = \frac{1}{\mathbf{v}_3^*}$ for Case 2.

Second, we use the previous result to show that $\mathbf{v} \succ \mathbf{v}'$. Partition Δ into three subsets, i.e. $\Delta = \Delta^{100'} \cup \Delta^{\hat{\mathbf{v}}} \cup \Delta^{001'}$ defined as $\Delta^{\hat{\mathbf{v}}} = \{\mathbf{v}'' \in \Delta \mid F(\mathbf{v}'') = F(\hat{\mathbf{v}})\}$, $\Delta^{100'} = \{\mathbf{v}'' \in \Delta \mid F(\mathbf{v}'') > F(\hat{\mathbf{v}})\}$ and $\Delta^{001'} = \{\mathbf{v}'' \in \Delta \mid F(\mathbf{v}'') < F(\hat{\mathbf{v}})\}$. We have by construction that $\mathbf{v} \in \Delta^{100'}$ and $\mathbf{v}' \in \Delta^{001'}$. We can show that $\mathbf{v}' \prec \hat{\mathbf{v}}$ using the same proof technique as above, i.e. show that $\mathbf{v}'$ is on a segment connecting $\hat{\mathbf{v}}$ to a $\mathbf{v}''$ on the edge E_{001}^{100} with $\mathbf{v}_1'' < \hat{\mathbf{v}}_1^*$ and, hence, such that $\mathbf{v}'' \prec \hat{\mathbf{v}}$. By the convexity property of $\succeq$, this yields in turn $\mathbf{v}' \prec \hat{\mathbf{v}}$. Similarly, we can show that $\mathbf{v} \succ \hat{\mathbf{v}}$ by showing that $\mathbf{v}$ is on a segment connecting $\hat{\mathbf{v}}$ to a $\mathbf{v}'''$ that is *either* on the edge E_{001}^{100} with $\mathbf{v}_1''' > \hat{\mathbf{v}}_1^*$ and, hence, such that $\mathbf{v}''' \succ \hat{\mathbf{v}}$ *or* on the edge E_{010}^{100} and, as $\mathbf{v}^{100} \succ \mathbf{v}^{010} \succ \hat{\mathbf{v}}$, such that $\mathbf{v}''' \succ \hat{\mathbf{v}}$. This implies in both cases that $\mathbf{v} \succ \hat{\mathbf{v}}$.

2.2 Proof of Proposition 2

We first prove necessity. Proving that GD satisfies Independence of Dead* is straightforward and left to the reader. Proposition 1 (see Appendix 2.3) shows that GD satisfies ID Equivalence. Finally, GD satisfies Additive Decomposibility when the size function is defined as $\eta(\mathbf{x}, \mu) = f(\mathbf{x}) + p(\mathbf{x}) + d^{GD}(\mathbf{x}, \mu)$. We show that this function is indeed such that $\eta(\mathbf{x}, \mu) = \eta(\mathbf{x}', \mu') + \eta(\mathbf{x}'', \mu'')$. Given that $f(\mathbf{x}', \mathbf{x}'') + p(\mathbf{x}', \mathbf{x}'') = f(\mathbf{x}') + p(\mathbf{x}') + f(\mathbf{x}'') + p(\mathbf{x}'')$, there remains to show that $d^{GD}((\mathbf{x}', \mathbf{x}''), \mu) = d^{GD}(\mathbf{x}', \mu') + d^{GD}(\mathbf{x}'', \mu'')$.

We have

$$d^{GD}((\mathbf{x}', \mathbf{x}''), \mu) = \sum_{a=0}^{\hat{a}-1} n_a(\mathbf{x}', \mathbf{x}'') * \mu_a * (\hat{a} - (a+1))$$
$$= \sum_{a=0}^{\hat{a}-1} (n_a(\mathbf{x}') + n_a(\mathbf{x}'')) * \frac{n_a(\mathbf{x}') * \mu'_a + n_a(\mathbf{x}'') * \mu''_a}{n_a(\mathbf{x}') + n_a(\mathbf{x}'')} * (\hat{a} - (a+1))$$
$$= d^{GD}(\mathbf{x}', \mu') + d^{GD}(\mathbf{x}'', \mu'').$$

It is then straighforwd to verify Equation (4) by replacing $\mathbf{P}$ and η by their expressions.

We now prove sufficiency. Take any pair $(\mathbf{x}', \mu) \in \mathcal{O}$. Consider the distribution $\mathbf{x}$ obtained from $\mathbf{x}'$ by removing all dead individuals in $\mathbf{x}$. We have $\mathbf{P}(\mathbf{x}, \mu) = \mathbf{P}(\mathbf{x}', \mu)$ by Independence of Dead* and also $GD_\gamma(\mathbf{x}, \mu) = GD_\gamma(\mathbf{x}', \mu)$.

The proof requires to define, for each $a \in \{0, \ldots, a^*\}$, two counterfactual pairs $(\mathbf{x}^*_a, \mu^*_a)$ and $(\mathbf{x}^0_a, \mu^0_a)$, which are illustrated in the center and right panels of Figure 2.

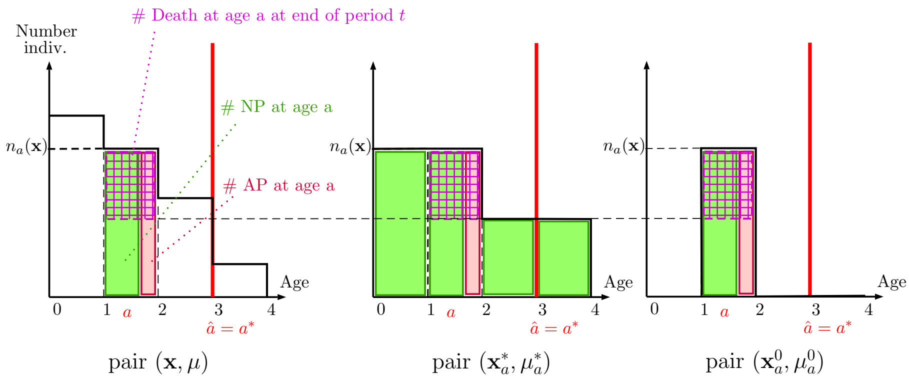

Figure 2: Left panel: pair $(\mathbf{x}, \mu)$. Center panel: stationary pair $(\mathbf{x}^*_a, \mu^*_a)$, where dead individuals are not shown. Right panel: degenerated pair $(\mathbf{x}^0_a, \mu^0_a)$.

The counterfactual pair $(\mathbf{x}^*_a, \mu^*_a)$ is stationary. The vector μ^*_a is such that mortality rates are zero except for two cases: $\mu^*_a = \mu_a$ and $\mu^*_{a^*} = 1$, which is $\mu^*_a = (0, \ldots, 0, \mu_a, 0, \ldots, 0, 1)$. We now turn to the construction of the distribution $\mathbf{x}^*_a$. At all ages $a' \leq a$, there are exactly $n_a(\mathbf{x})$ alive individuals (i.e. $n_{a'}(\mathbf{x}^*_a) = n_a(\mathbf{x})$); for all ages $a' > a$ we have $n_{a'}(\mathbf{x}^*_a) = n_a(\mathbf{x}) * (1 - \mu_a)$. At all ages $a' \leq a$, there are no dead individuals; for all ages $a' > a$, this number is $n_a(\mathbf{x}) * \mu_a$. There are no poor individuals in $\mathbf{x}_a^*$ except at age a, where this number is equal to the number of a-years old individuals in $\mathbf{x}$ whose status is AP, i.e. $\#\{i \leq n(\mathbf{x}) | s_i = AP \text{ and } b_i = t - a\}$.

The counterfactual "degenerated" pair $(\mathbf{x}_a^0, \mu_a^0)$ is not stationary and all its alive individuals are a-years old. The vector $\mu_a^0 = \mu_a^*$, which is $\mu_a^0 = (0, \ldots, 0, \mu_a, 0, \ldots, 0, 1)$. We now turn to the construction of distribution $\mathbf{x}_a^0$. At all ages $a' \neq a$, there are no alive individuals (i.e. $n_{a'}(\mathbf{x}_a^0) = 0$); and we have $n_a(\mathbf{x}_a^0) = n_a(\mathbf{x})$. There are no dead individuals. The number of AP individuals in $\mathbf{x}_a^0$ is equal to the number of a-years old individuals in $\mathbf{x}$ whose status is AP, i.e. $\#\{i \leq n(\mathbf{x}) | s_i = AP \text{ and } b_i = t - a\}$.

By iterative application of Additive Decomposibility, we have that

$$\mathbf{P}(\mathbf{x}, \mu) = \frac{\sum_{j=0}^{a^*} \eta(\mathbf{x}_j^0, \mu_j^0) * \mathbf{P}(\mathbf{x}_j^0, \mu_j^0)}{\sum_{j=0}^{a^*} \eta(\mathbf{x}_j^0, \mu_j^0)}. \tag{2}$$

By Independence of Dead*, Equation (2) also holds for the stationary distribution $(\mathbf{x}_a^*, \mu_a^*)$:

$$\mathbf{P}(\mathbf{x}_a^*, \mu_a^*) = \frac{\sum_{j=0}^{a^*} \eta((\mathbf{x}_a^*)_j^0, (\mu_a^*)_j^0) * \mathbf{P}((\mathbf{x}_a^*)_j^0, (\mu_a^*)_j^0)}{\sum_{j=0}^{a^*} \eta((\mathbf{x}_a^*)_j^0, (\mu_a^*)_j^0)}, \tag{3}$$

where the pair $((\mathbf{x}_a^*)_j^0, (\mu_a^*)_j^0)$ is the degenerated pair for age j associated to the stationary pair $(\mathbf{x}_a^*, \mu_a^*)$, i.e. the mortality vector $(\mu_a^*)_j^0 = \mu_a^*$ for $j = a$ and $(\mu_a^*)_j^0 = (0, \ldots, 0, 1)$ for $j \neq a$; and for $j = a$ we have $((\mathbf{x}_a^*)_j^0, (\mu_a^*)_j^0) = (\mathbf{x}_a^0, \mu_a^0)$.

For all $j \neq a$ we show that $\mathbf{P}((\mathbf{x}_a^*)_j^0, (\mu_a^*)_j^0) = 0$. Recall that $(\mu_a^*)_j^0 = (0, \ldots, 0, 1)$ and that distribution $(\mathbf{x}_a^*)_j^0$ has no individual whose status is AP. Consider the stationary pair $(\mathbf{x}''', \mu''')$ such that $\mu''' = (0, \ldots, 0, 1)$, distribution $\mathbf{x}'''$ has no individual whose status is AP and no dead individual. Provided that $n(\mathbf{x}''') = n_j(\mathbf{x}_a^*) * (a^* + 1)$, we have that $\mathbf{P}((\mathbf{x}_a^*)_j^0, (\mu_a^*)_j^0)$ appears in the decomposition (2) applied to $(\mathbf{x}''', \mu''')$. By construction, $ID_\gamma(\mathbf{x}''', \mu''') = 0$. By ID Equivalence, we have that $\mathbf{P}(\mathbf{x}''', \mu''') = ID_\gamma(\mathbf{x}''', \mu''') = 0$. Given that $\mathbf{P}$ does not yield negative images, we must have that $\mathbf{P}((\mathbf{x}_a^*)_j^0, (\mu_a^*)_j^0) = 0$.[10]

As $(\mathbf{x}_a^*, \mu_a^*)$ is stationary, we have from ID Equivalence that $\mathbf{P}(\mathbf{x}_a^*, \mu_a^*) = ID_\gamma(\mathbf{x}_a^*, \mu_a^*)$ for some $\gamma > 0$. As GD satisfies ID Equivalence, we have $\mathbf{P}(\mathbf{x}_a^*, \mu_a^*) = GD_\gamma(\mathbf{x}_a^*, \mu_a^*)$. Given that $\mathbf{P}((\mathbf{x}_a^*)_j^0, (\mu_a^*)_j^0) = 0$ for all $j \neq a$, and $\sum_{j=0}^{a^*} \eta((\mathbf{x}_a^*)_j^0, (\mu_a^*)_j^0) = \eta(\mathbf{x}_a^*, \mu_a^*)$,

[10] To be complete, there remains to show that $\eta((\mathbf{x}_a^*)_j^0, (\mu_a^*)_j^0) > 0$ when $n_j(\mathbf{x}_a^*) > 0$. If it is not the case, one can derive a contradiction with the requirement that $\eta(\mathbf{x}, \mu) = \eta(\mathbf{x}', \mu') + \eta(\mathbf{x}'', \mu'')$.

Equation (3) may be rewritten as

$$\mathbf{P}((\mathbf{x}_a^*)_a^0, (\mu_a^*)_a^0) = \frac{\eta(\mathbf{x}_a^*, \mu_a^*) * GD_\gamma(\mathbf{x}_a^*, \mu_a^*)}{\eta((\mathbf{x}_a^*)_a^0, (\mu_a^*)_a^0)}.$$

As $((\mathbf{x}_a^*)_a^0, (\mu_a^*)_a^0) = (\mathbf{x}_a^0, \mu_a^0)$, this last identity becomes

$$\mathbf{P}(\mathbf{x}_a^0, \mu_a^0) = \frac{\eta(\mathbf{x}_a^*, \mu_a^*) * GD_\gamma(\mathbf{x}_a^*, \mu_a^*)}{\eta(\mathbf{x}_a^0, \mu_a^0)}.$$

Inserting this last expression in Equation (2), where $\sum_{j=0}^{a^*} \eta(\mathbf{x}_j^0, \mu_j^0) = \eta(\mathbf{x}, \mu)$, yields

$$\mathbf{P}(\mathbf{x}, \mu) = \frac{\sum_{j=0}^{a^*} \eta(\mathbf{x}_j^*, \mu_j^*) * GD_\gamma(\mathbf{x}_j^*, \mu_j^*)}{\eta(\mathbf{x}, \mu)}. \tag{4}$$

Equation (4) holds for all pairs. If we have that function η is defined as $\eta(\mathbf{x}, \mu) = f(\mathbf{x}) + p(\mathbf{x}) + d^{GD}(\mathbf{x}, \mu)$, then Equation (4) simplifies to $\mathbf{P}(\mathbf{x}, \mu) = GD_\gamma(\mathbf{x}, \mu)$ and the proof is complete. We now show that the function η is indeed expressed as $\eta(\mathbf{x}, \mu) = f(\mathbf{x}) + p(\mathbf{x}) + d^{GD}(\mathbf{x}, \mu)$. Equation (4) holds in particular for any stationary pair $(\mathbf{x}', \mu')$. Therefore, by ID Equivalence we have

$$GD_\gamma(\mathbf{x}', \mu') = \frac{\sum_{j=0}^{a^*} \eta((\mathbf{x}')_j^*, (\mu')_j^*) * GD_\gamma((\mathbf{x}')_j^*, (\mu')_j^*)}{\eta(\mathbf{x}', \mu')}, \tag{5}$$

which only holds if function η has the appropriate expression.

2.3 Proof that GD satisfies ID Equivalence

Lemma 1 implies that GD satisfies ID Equivalence, but is slightly more general. This lemma shows that GD is equivalent to ID as soon as natality is constant for the $\hat{a} - 1$ periods preceding t and mortality rates are consistent with the population pyramid up to age $\hat{a} - 1$.

Lemma 1 (Equivalence between GD and ID indices in stationary pairs)**.** *If the pair $(\mathbf{x}, \mu) \in \mathcal{O}$ is such that for some $n^* \in \mathcal{N}$ and all $a \in \{0, \ldots, \hat{a} - 1\}$, we have:*

- $n_a(\mathbf{x}) + d_a(\mathbf{x}) = n^* \in \mathcal{N}$,
- $n_{a+1}(\mathbf{x}) = n_a(\mathbf{x}) * (1 - \mu_a)$,

then we have that $GD_\gamma(\mathbf{x}, \mu) = ID_\gamma(\mathbf{x})$.

Proof. By the definitions of GD_γ and ID_γ, we have $GD_\gamma(\mathbf{x}, \mu) = ID_\gamma(\mathbf{x})$ if $d(\mathbf{x}) = d^{GD}(\mathbf{x}, \mu)$. There remains to show that $d(\mathbf{x}) = d^{GD}(\mathbf{x}, \mu)$.

By definition, the number of prematurely dead individuals counted in period t by the inherited deprivation approach is

$$d(\mathbf{x}) = \sum_{a=1}^{\hat{a}-1} d_a(\mathbf{x}).$$

As the number of newborns is assumed constant in the $\hat{a} - 1$ periods preceding t,

$$d(\mathbf{x}) = \sum_{a=1}^{\hat{a}-1} \left(n^* - n_a(\mathbf{x})\right).$$

As $n_0(\mathbf{x}) = n^*$, we may rewritte the previous equation as

$$d(\mathbf{x}) = \sum_{a=1}^{\hat{a}-1} \left(\sum_{a'=0}^{a-1} (n_{a'}(\mathbf{x}) - n_{a'+1}(\mathbf{x})) \right),$$

and developing the sums, we get

$$\begin{aligned} d(\mathbf{x}) &= (\hat{a} - 1)(n_0(\mathbf{x}) - n_1(\mathbf{x})) + (\hat{a} - 2)(n_1(\mathbf{x}) - n_2(\mathbf{x})) + \cdots + (\hat{a} - (\hat{a} - 1))(n_{\hat{a}-2}(\mathbf{x}) - n_{\hat{a}-1}(\mathbf{x})), \\ &= \sum_{a=0}^{\hat{a}-2} (n_a(\mathbf{x}) - n_{a+1}(\mathbf{x}))(\hat{a} - (a + 1)), \end{aligned}$$

and given that $\hat{a} - ((\hat{a} - 1) + 1) = 0$, this is equivalent to

$$d(\mathbf{x}) = \sum_{a=0}^{\hat{a}-\mathbf{1}} (n_a(\mathbf{x}) - n_{a+1}(\mathbf{x}))(\hat{a} - (a + 1)).$$

Finally, as $n_{a+1}(\mathbf{x}) = n_a(\mathbf{x}) - n_a(\mathbf{x}) * \mu_a$ for all $a \in \{0, \ldots, \hat{a} - 2\}$, we get

$$d(\mathbf{x}) = \sum_{a=0}^{\hat{a}-1} n_a(\mathbf{x}) * \mu_a * (\hat{a} - (a + 1)) = d^{GD}(\mathbf{x}, \mu).$$

□

2.4 Proof of Proposition 3

We prove sufficiency. If GD_γ satisfies Monotonicity in Current Mortality, then for any two pairs $(\mathbf{x}, \mu), (\mathbf{x}, \mu') \in \mathcal{O}$ with $\mu_a \geq \mu'_a$ for all $a \in \{0, \ldots, \hat{a}-2\}$ we have $GD_\gamma(\mathbf{x}, \mu) \geq GD_\gamma(\mathbf{x}, \mu')$. As the two pairs share the same distribution $\mathbf{x}$, the precondition that $\mu_a \geq \mu'_a$ for all $a \in \{0, \ldots, \hat{a}-2\}$ implies $d^{GD}(\mathbf{x}, \mu) \geq d^{GD}(\mathbf{x}, \mu')$. As $GD_\gamma(\mathbf{x}, \mu) \geq GD_\gamma(\mathbf{x}', \mu)$, inequality $d^{GD}(\mathbf{x}, \mu) \geq d^{GD}(\mathbf{x}, \mu')$ implies that $\frac{\partial GD_\gamma}{\partial d^{GD}} \geq 0$. By chain derivation, we have

$$\frac{\partial GD_\gamma}{\partial d^{GD}} = \frac{\gamma(p+f) - p}{(p+f+d^{GD})^2}.$$

Thus, we have $\frac{\partial GD_\gamma}{\partial d^{GD}} \geq 0$ if and only if $\gamma \geq \frac{p}{p+f}$. As this must hold for all $\mathbf{x}$, we must have $\gamma \geq 1$.

The proof for necessity is easily obtained by reversing the above argument.

3 GD and ID count the same number of PYPLs

In the absence of migration, many aspects of distribution $\mathbf{x}$ are mechanically related to the distribution and mortality vector of the preceding period. We say that a distribution $\mathbf{x}'$ in period t is **generated** by the pair $(\mathbf{x}, \mu)$ in period $t-1$ if (i) the number of individuals born in each period before t is the same in both distributions, (ii) all individuals in $\mathbf{x}'$ who do not have a counterpart in $\mathbf{x}$ are newborns, (iii) individuals that are dead in $\mathbf{x}$ have their counterpart also dead in $\mathbf{x}'$ and (iv) the number of a-year-old individuals in $\mathbf{x}'$ is equal to the number of $(a-1)$-year-old individuals in $\mathbf{x}$ multiplited by the survival rate $1 - \mu_{a-1}$. Formally, that is

(i) $n_{a+1}(\mathbf{x}') + d_{a+1}(\mathbf{x}') = n_a(\mathbf{x}) + d_a(\mathbf{x})$ for all $a \geq 0$,

(ii) $b_j = t$ for all j present in $\mathbf{x}'$ but not in $\mathbf{x}$,

(iii) $(s'_i, b'_i) = (s_i, b_i)$ for all i present in $\mathbf{x}$ such that $s_i = D$,

(iv) $n_{a+1}(\mathbf{x}') = n_a(\mathbf{x}) * (1 - \mu_a)$ for all $a \geq 0$.

Let the set of all periods up to t be denoted by $\mathcal{Z}_t = \{-\infty, \ldots, t-1, t\}$. In theory, a pair $\left(\mathbf{x}^t, \mu^t\right)$ is the last element of a stream of pairs $(\mathbf{x}^\tau, \mu^\tau)_{\tau \in \mathcal{Z}_t} \in \mathcal{O}^{\mathcal{Z}_t}$, which are such that $\mathbf{x}^{\tau+1}$ is generated by $(\mathbf{x}^\tau, \mu^\tau)$ for all $\tau \in \mathcal{Z}_t$. When evaluating $(\mathbf{x}^\tau, \mu^\tau)_{\tau \in \mathcal{Z}_t}$ in time t, ID only considers the information in current alive population $\mathbf{x}^t$ as well as past natality $n_0(\mathbf{x}^\tau)$ for $\tau \in \{t-(\hat{a}-1), \ldots, t-1\}$. In turn, GD only consider the information in current alive population $\mathbf{x}^t$ together with current mortality μ^t.

Proposition 1 shows that GD and ID count the same number of PYPLs even outside stationary populations. Consider a population in a stationary state up to period 0. Assume that, over the time-frame $\{0, \ldots, t^*\}$, this population is hit by a series of mortality and natality shocks. From period $t^* + 1$ onwards, natality and mortality return to the values they took before period 0. Mechanically, the young part of the population pyramid may need up to $\hat{a} - 1$ periods after period t^* in order to come back to its previous stationary state. By Lemma 1, GD and ID indices are equal outside the (extended) time-frame $\{0, \ldots, t^* + \hat{a} - 1\}$. Proposition 1 shows that GD and ID compute the same number of PYPLs over the (extended) time frame. Let $n(\mathbf{x}^t) = p(\mathbf{x}^t) + f(\mathbf{x}^t)$ be the number of alive individuals in $\mathbf{x}^t$.

Proposition 1 (GD and ID count the same number of PYPLs)**.**
Let the stream of pairs $\left(\mathbf{x}^t, \mu^t\right)_{t \in \mathcal{Z}} \in \mathcal{O}^{\mathcal{Z}}$ *be such that* $\mathbf{x}^{t+1}$ *is generated by* $\left(\mathbf{x}^t, \mu^t\right)$ *for all* $t \in \mathcal{Z}$. *Take any* $t^* \geq 0$. *If we have for all* $t \in \mathcal{Z} \backslash_{\{0, \ldots, t^* + \hat{a} - 1\}}$ *that*

- $n_0(\mathbf{x}^t) = n^* \in \mathcal{N}_{++}$ *and* $\mu^t = \mu^* \in \mathcal{M}$

then we have

$$\sum_{t=0}^{t^*+\hat{a}-1} \big(n(\mathbf{x}^t) + d(\mathbf{x}^t)\big) * ID_\gamma(\mathbf{x}^t) = \sum_{t=0}^{t^*+\hat{a}-1} \big(n(\mathbf{x}^t) + d^{GD}(\mathbf{x}^t, \mu^t)\big) * GD_\gamma(\mathbf{x}^t, \mu^t), \qquad (6)$$

and for all $t \in \mathcal{Z} \backslash_{\{0, \ldots, t^* + \hat{a} - 1\}}$ *we have* $ID_\gamma(\mathbf{x}^t) = GD_\gamma(\mathbf{x}^t, \mu^t)$.

Proof. We prove the two implications in turn.

- **Step 1**: For all $t \in \mathcal{Z} \backslash_{\{0, \ldots, t^* + \hat{a} - 1\}}$ we have $ID_\gamma(\mathbf{x}^t) = GD_\gamma(\mathbf{x}^t, \mu^t)$.

  Take any $t \in \mathcal{Z} \backslash_{\{0, \ldots, t^* + \hat{a} - 1\}}$. By definition, we have $ID_\gamma(\mathbf{x}^t) = GD_\gamma(\mathbf{x}^t, \mu^t)$ if $d(\mathbf{x}^t) = d^{GD}(\mathbf{x}^t, \mu^t)$. By the proof of Lemma 1, we have $d(\mathbf{x}^t) = d^{GD}(\mathbf{x}^t, \mu^t)$ if (A) for all $a \in \{0, \ldots, \hat{a} - 2\}$ we have $n_{a+1}(\mathbf{x}^t) = n_a(\mathbf{x}^t) * (1 - \mu_a^t)$ and (B) for all cohorts $a \in \{0, \ldots, \hat{a} - 1\}$ we have $n_a(\mathbf{x}^t) + d_a(\mathbf{x}^t) = n^* \in \mathcal{N}_{++}$.

  First, we prove (A). As $t \notin \{0, \ldots, t^* + \hat{a} - 1\}$, we have by assumption that for all previous periods $t' \in \{t - (\hat{a} - 1), \ldots, t - 1\}$

  - distribution $\mathbf{x}^{t'+1}$ is generated by $(\mathbf{x}^{t'}, \mu^{t'})$,
  - $n_0(\mathbf{x}^{t'}) = n^*$ and $\mu^{t'} = \mu^*$.

As a result, we have for all $a \in \{0, \ldots, \hat{a}-2\}$ that

$$n_{a+1}(\mathbf{x}^t) = n^* * \prod_{l=0}^{a} (1 - \mu_l^*),$$

which implies for all $a \in \{1, \ldots, \hat{a}-2\}$ that $n_{a+1}(\mathbf{x}^t) = n_a(\mathbf{x}^t) * (1-\mu_a^*)$. Finally, last equation also holds for $a = 0$ as by $n_0(\mathbf{x}^t) = n^*$ we also have $n_1(\mathbf{x}^t) = n^* * (1-\mu_0^*)$, the desired result.

Second, we prove (B). Given that for all $t' \in \{t-(\hat{a}-1), \ldots, t-1\}$ distribution $\mathbf{x}^{t'+1}$ is generated by $(\mathbf{x}^{t'}, \mu^{t'})$, we have for all $a \in \{1, \ldots, \hat{a}-1\}$ that $n_a(\mathbf{x}^t) + d_a(\mathbf{x}^t) = n_0(\mathbf{x}^{t-a})$. As for all $t' \in \{t-(\hat{a}-1), \ldots, t\}$ we have $n_0(\mathbf{x}^{t'}) = n^*$, the previous claim implies that $n_a(\mathbf{x}^t) + d_a(\mathbf{x}^t) = n^*$ for all $a \in \{0, \ldots, \hat{a}-1\}$.

- **Step 2**: Equation (6) holds.

  By definition of $ID_\gamma(\mathbf{x}^t)$ and $GD_\gamma(\mathbf{x}^t, \mu^t)$, Equation (6) is

$$\sum_{t=0}^{t^*+\hat{a}-1} \big(n(\mathbf{x}^t) + d(\mathbf{x}^t)\big) * \left(\frac{n(\mathbf{x}^t) * HC(\mathbf{x}^t)}{n(\mathbf{x}^t) + d(\mathbf{x}^t)} + \gamma \frac{d(\mathbf{x}^t)}{n(\mathbf{x}^t) + d(\mathbf{x}^t)}\right) =$$
$$\sum_{t=0}^{t^*+\hat{a}-1} \big(n(\mathbf{x}^t) + d^{GD}(\mathbf{x}^t, \mu^t)\big) * \left(\frac{n(\mathbf{x}^t) * HC(\mathbf{x}^t)}{n(\mathbf{x}^t) + d^{GD}(\mathbf{x}^t, \mu^t)} + \gamma \frac{d^{GD}(\mathbf{x}^t, \mu^t)}{n(\mathbf{x}^t) + d^{GD}(\mathbf{x}^t, \mu^t)}\right),$$

  which is equivalent to

$$\underbrace{\sum_{t=0}^{t^*+\hat{a}-1} d(\mathbf{x}^t)}_{\equiv SID} = \underbrace{\sum_{t=0}^{t^*+\hat{a}-1} d^{GD}(\mathbf{x}^t, \mu^t)}_{\equiv SGD}. \tag{7}$$

  In order to prove Equation (7), we develop the sums SID and SGD.

  First, we develop SGD. Using the short notation $\Delta_a = \hat{a} - (a+1)$, the definition of d^{GD} is

$$d^{GD}(\mathbf{x}^t, \mu^t) = \sum_{a=0}^{\hat{a}-1} n_a(\mathbf{x}^t) * \mu_a^t * \Delta_a,$$

and as $\Delta_{\hat{a}-1} = 0$, we have

$$SGD = \sum_{t=0}^{t^*+\hat{a}-1} \sum_{a=0}^{\hat{a}-\mathbf{2}} n_a(\mathbf{x}^t) * \mu_a^t * \Delta_a. \tag{8}$$

Equation (8) shows that SGD counts the number of person-years prematurely lost (PYPL) due to deaths occuring in the time-frame $T = \{0, \dots, t^* + \hat{a} - 1\}$. All these PYPLs are lost for periods in $\{1, \dots, t^* + 2\hat{a} - 2\}$. Equation (9) divides these PYPLs between the PYPLs that are lost for periods in the time-frame T – i.e. for periods 1 to $t^* + \hat{a} - 1$ – and those lost for periods following the time-frame T – i.e. for periods $t^* + \hat{a}$ to $t^* + 2\hat{a} - 2$.

$$SGD' = \underbrace{\sum_{t=0}^{t^*+\hat{a}-2} \sum_{a=0}^{(\hat{a}-2)} \sum_{\tau=t+1}^{\min\{t+(\hat{a}-1)-a, t^*+\hat{a}-1\}} n_a(\mathbf{x}^t) * \mu_a^t}_{SGD'-inside\ T} + \underbrace{\sum_{t=t^*+1}^{t^*+\hat{a}-1} \sum_{a=0}^{(\hat{a}-2)} \sum_{\tau=t^*+\hat{a}}^{t+(\hat{a}-1)-a} n_a(\mathbf{x}^t) * \mu_a^t}_{SGD'-outside\ T}. \tag{9}$$

We show that $SGD' = SGD$. As term $n_a(\mathbf{x}^t) * \mu_a^t$ is independent on τ we have

$$SGD' = \sum_{t=0}^{t^*+\hat{a}-2} \sum_{a=0}^{(\hat{a}-2)} n_a(\mathbf{x}^t) * \mu_a^t \sum_{\tau=t+1}^{\min\{t+(\hat{a}-1)-a, t^*+\hat{a}-1\}} 1 + \sum_{t=t^*+1}^{t^*+\hat{a}-1} \sum_{a=0}^{(\hat{a}-2)} n_a(\mathbf{x}^t) * \mu_a^t \sum_{\tau=t^*+\hat{a}}^{t+(\hat{a}-1)-a} 1.$$

We consider the expression of SGD' in turn for the set of periods $\{0, \dots, t^*\}$, then for the set of periods $\{t^*+1, \dots, t^*+\hat{a}-2\}$ and finally for period $t^*+\hat{a}-1$. For each of these three sets of periods, we show that the expression of SGD' corresponds to the expression of SGD.

– Periods $t \in \{0, \dots, t^*\}$.
  We have for all $a \in \{0, \dots, \hat{a} - 2\}$ that

$$\min\{t + (\hat{a} - 1) - a, t^* + \hat{a} - 1\} = t + \hat{a} - 1 - a,$$

  which implies

$$\sum_{\tau=t+1}^{t+(\hat{a}-1)-a} 1 = \hat{a} - (a + 1) = \Delta_a,$$

and therefore

$$\sum_{t=0}^{t^*}\sum_{a=0}^{(\hat{a}-2)} n_a(\mathbf{x}^t) * \mu_a^t \sum_{\tau=t+1}^{\min\{t+(\hat{a}-1)-a, t^*+\hat{a}-1\}} 1 = \sum_{t=0}^{t^*}\sum_{a=0}^{(\hat{a}-2)} n_a(\mathbf{x}^t) * \mu_a^t * \Delta_a.$$

– Periods $t \in \{t^*+1, \dots, t^*+\hat{a}-2\}$.

We have for all $a \in \{0, \dots, \hat{a}-2\}$ that

$$\begin{aligned}
&\sum_{t=t^*+1}^{t^*+\hat{a}-2}\sum_{a=0}^{(\hat{a}-2)} n_a(\mathbf{x}^t) * \mu_a^t \sum_{\tau=t+1}^{\min\{t+(\hat{a}-1)-a, t^*+\hat{a}-1\}} 1 + \sum_{t=t^*+1}^{t^*+\hat{a}-2}\sum_{a=0}^{(\hat{a}-2)} n_a(\mathbf{x}^t) * \mu_a^t \sum_{\tau=t^*+\hat{a}}^{t+(\hat{a}-1)-a} 1, \\
&= \sum_{t=t^*+1}^{t^*+\hat{a}-2}\sum_{a=0}^{(\hat{a}-2)} n_a(\mathbf{x}^t) * \mu_a^t \left(\sum_{\tau=t+1}^{\min\{t+(\hat{a}-1)-a, t^*+\hat{a}-1\}} 1 + \sum_{\tau=t^*+\hat{a}}^{t+(\hat{a}-1)-a} 1 \right), \\
&= \sum_{t=t^*+1}^{t^*+\hat{a}-2}\sum_{a=0}^{(\hat{a}-2)} n_a(\mathbf{x}^t) * \mu_a^t \sum_{\tau=t+1}^{t+(\hat{a}-1)-a} 1, \\
&= \sum_{t=t^*+1}^{t^*+\hat{a}-2}\sum_{a=0}^{(\hat{a}-2)} n_a(\mathbf{x}^t) * \mu_a^t * \Delta_a.
\end{aligned}$$

– Period $t = t^*+\hat{a}-1$.

We have for all $a \in \{0, \dots, \hat{a}-2\}$ that

$$\sum_{\tau=t^*+\hat{a}}^{t^*+2(\hat{a}-1)-a} 1 = \hat{a} - (a+1) = \Delta_a.$$

and therefore

$$\sum_{t=t^*+\hat{a}-1}^{t^*+\hat{a}-1}\sum_{a=0}^{(\hat{a}-2)} n_a(\mathbf{x}^t) * \mu_a^t \sum_{\tau=t^*+\hat{a}}^{t+(\hat{a}-1)-a} 1 = \sum_{t=t^*+\hat{a}-1}^{t^*+\hat{a}-1}\sum_{a=0}^{(\hat{a}-2)} n_a(\mathbf{x}^t) * \mu_a^t * \Delta_a.$$

Second, we develop SID. The sum SID counts the number of person-years that are prematurely lost for periods in the time-frame $T = \{0, \dots, t^*+\hat{a}-1\}$. As distribution $\mathbf{x}^{t+1}$ is generated by $(\mathbf{x}^t, \mu^t)$ for all $t \in \mathcal{Z}$, these PYPLs are lost due to deaths occuring in the set of periods $t \in \{-(\hat{a}-1), \dots, t^*+\hat{a}-2\}$. We express SID by counting all these PYPLs in the following way:

– run all periods t at which the occurrence of a death potentially generates a PYPL for a period in T,

– for each such period t, run all age-cohorts a whose death generates a PYPL for a period in T,

– count all periods in T that are prematurely lost due to a death occuring at age a in period t.

Using this way of counting, the sum SID is

$$SID = \sum_{t=-(\hat{a}-1)}^{t^*+\hat{a}-2} \sum_{a=0}^{(\hat{a}-2)} \sum_{\tau=\max\{0,t+1\}}^{\min\{t+(\hat{a}-1)-a,t^*+\hat{a}-1\}} n_a(\mathbf{x}^t) * \mu_a^t. \tag{10}$$

We illustrate the order in which (10) counts all the relevant PYPLs in Figure 3.

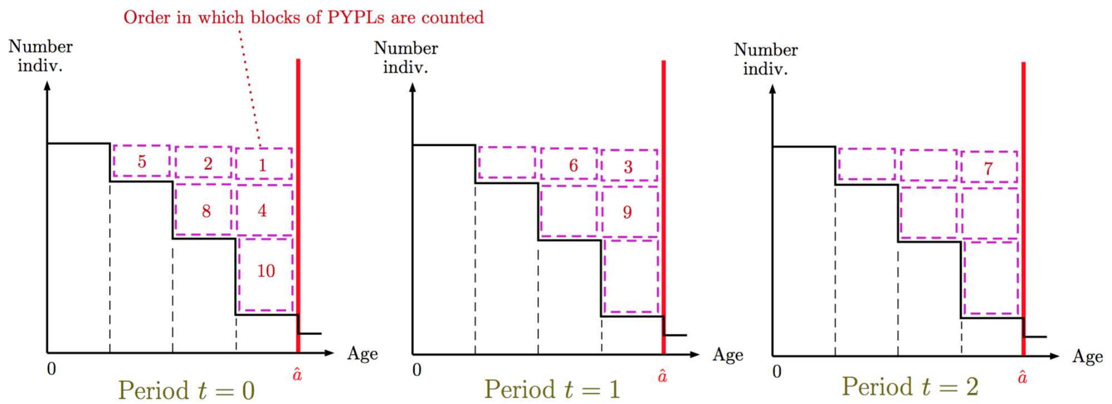

Figure 3: Order in which blocks of PYPLs are counted in (10). Block 1 corresponds to PYPLs due to newborns dying in period $-(\hat{a}-1)$. Block 2 and 3 are due to newborns dying in period $-(\hat{a}-2)$. Block 4 is due to 1-year-old dying in period $-(\hat{a}-2)$. Block 5, 6 and 7 are due to newborns dying in period $-(\hat{a}-3)$. Block 8 and 9 are due to 1-year-old dying in period $-(\hat{a}-3)$. Block 10 is due to 2-year-old dying in period $-(\hat{a}-3)$. For simplicity, this order is illustrated for pairs $(\mathbf{x}^t, \mu^t)$ that are stationary.

Equation (11) divides these PYPLs between the PYPLs generated by deaths occuring before the time-frame $T = \{0, \ldots, t^* + \hat{a} - 1\}$, and the PYPLs generated by deaths occuring during the time-frame T. This division yields the sum SID'

$$SID' = \underbrace{\sum_{t=-(\hat{a}-1)}^{-1} \sum_{a=0}^{t+(\hat{a}-1)} \sum_{\tau=0}^{t+(\hat{a}-1)-a} n_a(\mathbf{x}^t) * \mu_a^t}_{SID'-outside\ T} + \underbrace{\sum_{t=0}^{t^*+\hat{a}-2} \sum_{a=0}^{(\hat{a}-2)} \sum_{\tau=t+1}^{\min\{t+(\hat{a}-1)-a,t^*+\hat{a}-1\}} n_a(\mathbf{x}^t) * \mu_a^t}_{SID'-inside\ T}. \tag{11}$$

We show that $SID' = SID$. We consider the expression of SID' in turn for the set of periods $\{-(\hat{a}-1),\ldots,-1\}$ and then for the set of periods $\{0,\ldots,t^*+\hat{a}-2\}$. For each of these two sets of periods, we show that the expression of SID' corresponds to the expression of SID.

- Periods $t \in \{-(\hat{a}-1),\ldots,-1\}$.
  For this case we have $\max\{0, t+1\} = 0$. Furthermore, for all $a \geq 0$ we have $t - a \leq 0 \leq t^*$ and thus

$$\min\{t + (\hat{a}-1) - a, t^* + \hat{a} - 1\} = t + (\hat{a}-1) - a.$$

  Thus, we can rewritte the sum SID' on these periods as

$$\sum_{t=-(\hat{a}-1)}^{-1} \sum_{a=0}^{t+(\hat{a}-1)} \sum_{\tau=0}^{t+(\hat{a}-1)-a} n_a(\mathbf{x}^t) * \mu_a^t$$
$$= \sum_{t=-(\hat{a}-1)}^{-1} \sum_{a=0}^{t+(\hat{a}-1)} \sum_{\tau=\max\{0,t+1\}}^{\min\{t+(\hat{a}-1)-a,t^*+\hat{a}-1\}} n_a(\mathbf{x}^t) * \mu_a^t.$$

  Finally, for all $a > t + (\hat{a}-1)$ we have $t + (\hat{a}-1) - a < 0$, implying that

$$\sum_{\tau=0}^{t+(\hat{a}-1)-a} n_a(\mathbf{x}^t) * \mu_a^t = 0.$$

  and therefore the sum SID' is further rewritten as

$$\sum_{t=-(\hat{a}-1)}^{-1} \sum_{a=0}^{(\hat{a}-2)} \sum_{\tau=\max\{0,t+1\}}^{\min\{t+(\hat{a}-1)-a,t^*+\hat{a}-1\}} n_a(\mathbf{x}^t) * \mu_a^t.$$

- Periods $t \in \{0,\ldots,t^*+\hat{a}-2\}$.
  For this periods we have $\max\{0, t+1\} = t+1$. Hence, we can rewritte the sum SID' on these periods as

$$\sum_{t=0}^{t^*+\hat{a}-2} \sum_{a=0}^{(\hat{a}-2)} \sum_{\tau=t+1}^{\min\{t+(\hat{a}-1)-a,t^*+\hat{a}-1\}} n_a(\mathbf{x}^t) * \mu_a^t = \sum_{t=0}^{t^*+\hat{a}-2} \sum_{a=0}^{(\hat{a}-2)} \sum_{\tau=\max\{0,t+1\}}^{\min\{t+(\hat{a}-1)-a,t^*+\hat{a}-1\}} n_a(\mathbf{x}^t) * \mu_a^t.$$

Third, we show that $SGD' = SID'$. Given that SGD' *inside* T and SID' *inside* T are trivially equal to each other, we only need to show that SGD' *outside* $T =$

SID' *outside* T, which is

$$\sum_{t=t^*+1}^{t^*+\hat{a}-1} \sum_{a=0}^{(\hat{a}-2)} \sum_{\tau=t^*+\hat{a}}^{t+(\hat{a}-1)-a} n_a(\mathbf{x}^t) * \mu_a^t = \sum_{t=-(\hat{a}-1)}^{-1} \sum_{a=0}^{t+(\hat{a}-1)} \sum_{\tau=0}^{t+(\hat{a}-1)-a} n_a(\mathbf{x}^t) * \mu_a^t. \tag{12}$$

Part 1. We show that for both SGD' *outside* T and SID' *outside* T we have for all relevant t and a that

$$n_a(\mathbf{x}^t) * \mu_a^t = n_a^* * \mu_a^*,$$

where for all $a \in \{0, \ldots, \hat{a}-1\}$ we have $n_a^* = n^* \prod_{l=0}^{a-1}(1-\mu_l^*)$.

Recall that, by assumption, $\mathbf{x}^t$ is generated by $(\mathbf{x}^{t-1}, \mu^{t-1})$, $\mathbf{x}^{t-1}$ is generated by $(\mathbf{x}^{t-2}, \mu^{t-2})$, and so on. By assumption again, $n_0(\mathbf{x}^t) = n^* \in \mathcal{N}_{++}$ and $\mu^t = \mu^* \in \mathcal{M}$ for all $t \in \mathcal{Z}\backslash_{\{0,\ldots,t^*\}}$.

- SGD' *outside* T.

  Take any $t \in \{t^*+1, \ldots, t^*+\hat{a}-1\}$ and any $a \in \{0, \ldots, \hat{a}-2\}$ such that a death occuring in period t at age a generates a PYPL for a period outside the time-frame T, i.e. t and a are such that

  $$t + (\hat{a}-1) - a \geq t^* + \hat{a}.$$

  Given that $t \geq t^*+1$, we have that $\mu^t = \mu^*$ and hence $\mu_a^t = \mu_a^*$. There remains to show that

  $$n_a(\mathbf{x}^t) = n_0(\mathbf{x}^{t-a}) \prod_{l=0}^{a-1}(1-\mu_l^{t-a+l}) = n^* \prod_{l=0}^{a-1}(1-\mu_l^*) = n_a^*. \tag{13}$$

  As from period t^*+1 onwards, natality and mortality are fixed, last equation holds if $t-a \geq t^*+1$. This is the case for all t and a for which $t+(\hat{a}-1)-a \geq t^*+\hat{a}$, the desired result.

- SID' *outside* T.

  Take any $t \in \{-(\hat{a}-1), \ldots, -1\}$ and any $a \in \{0, \ldots, \hat{a}-2\}$. Given that $t \leq 0$, we have that $\mu_a^t = \mu_a^*$. Finally, given that natality and mortality are fixed for all periods preceding t, (13) holds, the desired result.

Part 2. We develop each side of Equation (12) using Part 1 and obtain identical mathematical expressions.

Consider first SGD' outside T. By Part 1, for all relevant t and a we have $n_a(\mathbf{x}^t) * \mu_a^t = n_a^* * \mu_a^*$. Therefore, we rewrite SGD' outside T as

$$\sum_{t=t^*+1}^{t^*+\hat{a}-1} \sum_{a=0}^{(\hat{a}-2)} \sum_{\tau=t^*+\hat{a}}^{t+(\hat{a}-1)-a} n_a(\mathbf{x}^t) * \mu_a^t = \sum_{t=t^*+1}^{t^*+\hat{a}-1} \sum_{a=0}^{(\hat{a}-2)} n_a^* * \mu_a^* \sum_{\tau=t^*+\hat{a}}^{t+(\hat{a}-1)-a} 1.$$

From last expression, the relevant t and a are such $t + (\hat{a} - 1) - a \geq t^* + \hat{a}$,[11] or yet $t \geq t^* + 1 + a$, hence

$$\begin{aligned}
&\sum_{t=t^*+1}^{t^*+\hat{a}-1} \sum_{a=0}^{(\hat{a}-2)} n_a^* * \mu_a^* \sum_{\tau=t^*+\hat{a}}^{t+(\hat{a}-1)-a} 1 \\
&\quad = \sum_{a=0}^{(\hat{a}-2)} \sum_{t=t^*+1+a}^{t^*+\hat{a}-1} n_a^* * \mu_a^* * (t - t^* - a), \\
&\quad = \sum_{a=0}^{(\hat{a}-2)} n_a^* * \mu_a^* * \left(\sum_{t=t^*+1+a}^{t^*+\hat{a}-1} t - (t^* + a) \sum_{t=t^*+1+a}^{t^*+\hat{a}-1} 1 \right), \\
&\quad = \sum_{a=0}^{(\hat{a}-2)} n_a^* * \mu_a^* * \left(\left((t^* + a) * (\hat{a} - 1 - a) + \sum_{t=1}^{\hat{a}-1-a} 1 \right) - (t^* + a)(\hat{a} - 1 - a) \right), \\
&\quad = \sum_{a=0}^{(\hat{a}-2)} n_a^* * \mu_a^* * \frac{(\hat{a} - 1 - a)(\hat{a} - a)}{2}.
\end{aligned}$$

Consider then SID' outside T. From Part 1 again, we may rewritte SID' outside T as

$$\sum_{t=-(\hat{a}-1)}^{-1} \sum_{a=0}^{(\hat{a}-2)} \sum_{\tau=0}^{t+(\hat{a}-1)-a} n_a(\mathbf{x}^t) * \mu_a^t = \sum_{t=-(\hat{a}-1)}^{-1} \sum_{a=0}^{(\hat{a}-2)} n_a^* * \mu_a^* \sum_{\tau=0}^{t+(\hat{a}-1)-a} 1.$$

From last expression, the relevant t and a are such $t + (\hat{a} - 1) - a \geq 0$,[12] or yet

[11] If $t + (\hat{a} - 1) - a < t^* + \hat{a}$, then we have $\sum_{\tau=t^*+\hat{a}}^{t+(\hat{a}-1)-a} 1 = 0$.

[12] If $t + (\hat{a} - 1) - a < 0$, then we have $\sum_{\tau=0}^{t+(\hat{a}-1)-a} 1 = 0$.

$t \geq a - (\hat{a} - 1)$. Therefore, we rewrite SID' outside T as

$$
\begin{aligned}
&\sum_{t=-(\hat{a}-1)}^{-1} \sum_{a=0}^{(\hat{a}-2)} n_a^* * \mu_a^* \sum_{\tau=0}^{t+(\hat{a}-1)-a} 1 \\
&\quad = \sum_{a=0}^{(\hat{a}-2)} \sum_{t=a-(\hat{a}-1)}^{-1} n_a^* * \mu_a^* * (t + \hat{a} - a) \\
&\quad = \sum_{a=0}^{(\hat{a}-2)} n_a^* * \mu_a^* * \left(\sum_{t=a-(\hat{a}-1)}^{-1} t + (\hat{a} - a) \sum_{t=a-(\hat{a}-1)}^{-1} 1 \right), \\
&\quad = \sum_{a=0}^{(\hat{a}-2)} n_a^* * \mu_a^* * \left(- \sum_{t'=1}^{\hat{a}-1-a} t' + (\hat{a} - a) * (\hat{a} - 1 - a) \right), \\
&\quad = \sum_{a=0}^{(\hat{a}-2)} n_a^* * \mu_a^* * \left(- \frac{(\hat{a} - 1 - a)(\hat{a} - a)}{2} + (\hat{a} - a) * (\hat{a} - 1 - a) \right), \\
&\quad = \sum_{a=0}^{(\hat{a}-2)} n_a^* * \mu_a^* * \frac{(\hat{a} - 1 - a)(\hat{a} - a)}{2},
\end{aligned}
$$

which shows that (12) holds.

□

4 List of countries in dataset

Table 2 lists all the countries and region of the dataset, along with their HC and GD as measured in 2015.

Table 2: Countries's deprivation measures in 2014

Country/Region	**HC**	**GD(50)**
World	11.7	17.3
East Asia Pacific	2.3	4.6
Europe Central Asia	1.7	4.5
Latin America Caribbean	4.3	7.6
Middle East North Africa	4.3	8.8
South Asia	12.4	18.3
Sub Saharan Africa	41.2	50.3

Table 2: Countries's deprivation measures in 2014

Country/Region	**HC**	**GD(50)**
China	0.7	2.3
Fiji	1.0	5.5
Indonesia	7.2	10.9
Lao PDR	17.7	25.8
Malaysia	0.0	2.1
Micronesia Fed Sts	15.1	19.1
Mongolia	0.3	5.9
Myanmar	6.2	12.1
Papua New Guinea	29.2	37.9
Philippines	7.8	13.1
Samoa	1.0	4.3
Solomon Islands	24.8	30.1
Thailand	0.0	2.2
Tonga	1.0	4.5
Vanuatu	15.3	21.5
Vietnam	2.4	4.9
Albania	0.7	2.4
Armenia	1.9	3.4
Azerbaijan	0.0	4.6
Belarus	0.0	1.6
Bosnia and Herzegovina	0.1	1.0
Bulgaria	1.2	2.4
Georgia	3.8	5.8
Kazakhstan	0.0	3.4
Kyrgyz Republic	2.5	6.7
Macedonia FYR	5.2	6.4
Moldova	0.0	2.3
Montenegro	0.0	0.9
Romania	5.7	7.0
Russian Federation	0.0	2.8
Serbia	0.1	1.0
Tajikistan	4.8	12.2

Table 2: Countries's deprivation measures in 2014

Country/Region	**HC**	**GD(50)**
Turkey	0.3	2.4
Turkmenistan	2.8	7.8
Ukraine	0.1	3.2
Uzbekistan	14.0	17.7
Belize	12.3	16.0
Bolivia	6.4	11.2
Brazil	3.4	6.5
Colombia	4.5	7.2
Costa Rica	1.5	3.3
Dominican Republic	1.8	6.6
Ecuador	3.4	6.8
El Salvador	1.9	6.7
Guatemala	7.9	13.5
Guyana	6.6	11.4
Haiti	23.5	32.1
Honduras	16.2	20.4
Jamaica	1.8	4.8
Mexico	3.4	6.2
Nicaragua	2.9	6.0
Paraguay	1.9	4.7
Peru	3.6	6.5
St Lucia	6.3	8.7
Suriname	18.8	22.4
Algeria	0.4	3.7
Djibouti	19.3	26.0
Egypt Arab Rep	1.3	5.9
Iran Islamic Rep	0.3	3.5
Iraq	2.4	8.3
Jordan	0.2	3.2
Lebanon	0.0	1.9
Morocco	0.9	4.6
Syrian Arab Rep	21.2	28.0
Tunisia	0.4	2.2

Table 2: Countries's deprivation measures in 2014

Country/Region	**HC**	**GD(50)**
Yemen Rep	30.4	36.9
Bangladesh	15.2	19.9
Bhutan	0.9	4.9
India	13.4	18.9
Nepal	7.0	12.4
Pakistan	5.3	15.8
Sri Lanka	0.7	2.4
Angola	28.2	38.9
Benin	49.5	57.8
Botswana	16.2	21.4
Burkina Faso	42.8	54.5
Burundi	74.8	78.8
Cabo Verde	7.2	11.1
Cameroon	22.8	33.4
Central African Rep	77.7	83.5
Chad	33.9	49.9
Comoros	18.1	24.2
Congo Dem Rep	71.7	76.7
Congo Rep	34.9	41.9
Cote d Ivoire	28.2	39.7
Ethiopia	30.9	39.3
Gabon	4.0	11.8
Gambia	11.0	20.3
Ghana	13.2	22.2
Guinea	32.8	45.5
Guinea Bissau	65.3	71.2
Kenya	37.3	43.5
Lesotho	54.8	62.4
Liberia	39.4	48.2
Madagascar	77.5	80.9
Malawi	70.2	74.7
Mali	47.7	60.0
Mauritania	6.2	14.0

Table 2: Countries's deprivation measures in 2014

Country/Region	**HC**	**GD(50)**
Mauritius	0.3	2.4
Mozambique	61.6	68.4
Namibia	13.4	20.9
Niger	44.2	56.5
Nigeria	47.0	57.5
Rwanda	55.2	59.7
Senegal	33.9	40.4
Sierra Leone	48.5	59.1
South Africa	18.9	26.2
Sudan	7.7	17.4
Tanzania	40.7	48.6
Togo	49.2	55.7
Uganda	39.4	47.9
Zambia	57.5	63.5
Zimbabwe	16.6	27.6